NORTH-HOLLAND

A. C. Aitken and the Consolidation of Matrix Theory

Richard William Farebrother
Victoria University of Manchester
Manchester, United Kingdom

Submitted by Simo Puntanen

ABSTRACT

We briefly outline the origins of formal matrix theory in the 1870s and discuss Aitken's role in the dissemination of matrix methods in the 1940s with particular reference to the subject area of statistics and economics. © 1997 Elsevier Science Inc.

1. INTRODUCTION

The year 1995 marks the one-hundreth anniversary of the birth of Alexander Craig Aitken (1895–1967). It therefore seems appropriate to begin this paper with a brief account of his life. For more detailed biographies see Aitken (1995), Tee (1980), and Whittaker (1968).

Alexander Craig Aitken was born in Dunedin, New Zealand, on 1 April 1895. He was educated at the University of Otago, Dunedin, New Zealand, where he was awarded the degree of M. A. in 1919, having spent the years 1915–1918 as an infantryman in ANZAC. After four years as a school teacher in Otago, he resumed his formal education at the University of Edinburgh, Edinburgh. Scotland, were he was awarded the degree of D.Sc. (in place of the anticipated Ph.D.,) in 1925. He joined the teaching staff of the University of Edinburgh, rising to the rank of Reader. On the retirement of Sir Edmund Taylor Whittaker (1873–1956) in 1946, Aitken was elected to the vacant chair of mathematics, a post he held until his own retirement in 1965. Aitken died in Edinburgh on 3 November 1967.

Although Aitken's name in nowadays perhaps best known to statisticians in connection with his (1935) work on the optimality properties of the

LINEAR ALGEBRA AND ITS APPLICATIONS 264:3–12 (1997)

© 1997 Elsevier Science Inc. All rights reserved.
655 Avenue of the Americas, New York, NY 10010
0024-3795/97/$17.00
PII S0024-3795(96)00398-9

generalized least squares estimator, he also deserves to be remembered by those concerned with the application of matrix techniques in this area as the sole author of the popular textbook *Determinants and Matrices* (1939a) and the joint author [with Herbert Westren Turnbull (1885–1961)] of the important monograph *An Introduction to the Theory of Canonical Matrices* (1932).

In the sixteenth and seventeenth centuries the word matrix was used to refer to the uterus or womb of a female creature. This usage is most familiar from Tyndale's translation of the Bible into the English language and in the subsequent edition authorized by King James I; for example, see the Book of Exodus, Chapter 13, verse 15 and Chapter 34, verse 19.

In the eighteenth century this usage was extended to geology, pottery, and foundry work. In geology the word matrix refers to the mass of rock in which a fossil or gemstone is embedded, and in pottery and foundrywork it refers to the mold in which the utensil is to be cast.

The modern mathematical use of the word was introduced by Sylvester in 1850 in the following terms (1850: 1904, p. 150):

> we must commence, not with a square, but with an oblong arrangement of terms consisting, suppose, of *m* lines and *n* columns. This will not in itself represent a determinant, but is as it were, a Matrix out of which we may form various systems of determinants...

whilst in the following year he notes (1851: 1904, p. 247):

> I have in previous papers defined a Matrix as a regular array of terms, out of which different systems of determinants may be engendered, as from the womb of a common parent...

This terminology was subsequently popularized in an important study by Cayley (1858).

The gradual emergence of a unified theory of matrices has been described by Hawkins (1975), and the dissemination of this theory by Grattan-Guinness and Ledermann (1994). In the present paper we will concentrate on the latter aspect of the subject and shall offer a detailed critique of a passage from Grattan-Guinness and Ledermann's account with particular reference to statistics and economics.

2. THE ORIGINS OF MATRIX THEORY

In an important paper on the history of linear algebra, Hawkins (1975) has argued that the traditional association of the name of Cayley with the crucial step in the foundation of formal matrix theory is overly generous to him, as it places too much emphasis on the surface features of the subject and too little on the underlying structure. Instead, Hawkins prefers to name Frobenius as the author of the paper most directly responsible for the foundation of matrix theory. In his paper, Frobenius (1877) unified the two main streams of ideas which contributed most to our present understanding of the subject, namely the gradual development of a spectral theory of functional equations by Lagrange (1775), Jacobi (1827), Cauchy (1829), Sturm (1829), and Weierstrass (1868), and the more rapid development of a symbolic algebra of substitutions (we would say transformations) associated with the names of Gauss (1801), Eisenstein (1844), Hermite (1854), and Cayley (1858).

This general conclusion regarding the origins of modern matrix theory in the 1870s was endorsed by Grattan-Guinness and Ledermann (1994) and Grattan-Guinness (1994) in their leading articles on the history of matrix theory and linear and nonlinear programming.

3. THE CONSOLIDATION OF MATRIX THEORY

Grattan-Guinnesss and Ledermann (1994, pp. 784–785) discussed the early textbooks on matrix theory in the following terms:

of these, the *Introduction to Higher Algebra* by the Harvard mathematician Maxime Bôcher (*1907*)... was an important pioneering work... Another noteworthy author was the Cambridge-trained mathematician C. E. Cullis, who disgorged his lectures at the University of Calcutta into three large volumes on *Matrices and Determinoids* (*1913–1925*). Much non-standard terminology reduced its influence, but it was notable in several ways: in particular, he stressed rectangular matrices from the start (indeed, a "determinoid" was his extension of a determinant to such matrices, using the Laplace expansion).

In passages quoted in the following section Anderson (1995) and Olkin (1995) confirm the central role of Bôchers's textbook in the dissemination of matrix theory in the U.S.A. in the 1930s and 1940s. Cullis's book was clearly less important than Bôchers's. However, in view of the statement at the end of the above quotation, it seems reasonable to suggest that a careful study of Cullis's work by a team of mathematicians concerned with the properties of the generalized inverses of matrices would pay dividends by offering insights into the nature of the subject and identifying possible sources of this and related concepts.

Returning to the passage from Grattan-Guinness and Ledermann (1994, pp. 784–785), we find that it continued:

Even then, though diffusion and education of matrix theory was slow; for example, when, in the mid-1920s, the creators of quantum mechanics were looking for techniques, matrix theory was still not widely known... The early 1930s saw several new textbooks, of which Turnbull and Aitken 1932 was popular.

The rise of matrix theory to staple diet has occurred only since the 1950s... Mirsky *1955* was one of the better books in English.

For more detailed assessment of Aitken's contribution to the popularization of matrix algebra we turn to Ledermann (1968, pp. 164–165):

With his flair for elegant formalism Aitken was quick to realize the usefulness of matrix algebra as a powerful tool in many branches of mathematics. At a time when matrix techniques were not yet widely known he applied matrix algebra with striking success to certain statistical problems.

His interest in matrices was shared by H. W. Turnbull. Their joint book *Canonical Matrices* soon became a standard work on the subject.

Ledermann (1968, p. 165) continues his analysis by describing the text of Aitken's (1939a) *Determinants and Matrices* in the following terms:

Charmingly written, it is typical of Aitken's style, and the choice of topics is characteristic of his personal taste and his attitude to algebra. It is quite unlike any of the numerous treatises on linear algebra, which have appeared since the 1930's. Although linear equations are treated in some detail, there is no explicit mention of vector spaces or linear mappings and only a few brief paragraphs are devoted to latent roots and vectors.

Although these remarks would seem to suggest that Aitken's book is no longer suitable as a text on matrix algebra, it is interesting to note that it is still in print some 56 years later, if only at price of £44.75. Further, Aitken seems to have thought the subject too advanced for statisticians, as, apart from a brief account of the solution of a set of normal equations by the method of Gaussian elimination, there does not seem to be any discussion of matrix methods in his companion text on statistical mathematics (1939b).

4. DISSEMINATION OF MATRIX THEORY IN STATISTICS

In his detailed comments on an earlier version of the present paper Watson (1995) remarked:

I can assure you that few *English* knew matrix algebra in 1950—and none seemed to see the geometrical background.... Yet my 1947 lecture notes from R C Bose use vector spaces. I didn't see the light until I read Halmos's [1942] book—and I feel this was the turning point for everyone. Maybe the translation of Courant & Hilbert [1938] helped too. Certainly the bloom of Functional analysis around 1950 helped....

I never heard of matrices as an undegraduate (c. 1941)—a German refugee introduced them to me in 1943 and another taught Alan James in Adelaide even better.

By contrast with Watson's first sentence, Plackett (1996) notes

that the courses on matrix algebra which I attended at Cambridge in 1940 or 1941 gave an excellent basis for my forays into matrix methods in statistics later in the decade.

Watson (1996) subsequently admitted that he should perhaps not have constructed so broad a generalization from his impression of the situation which faced him on his arrival in England in 1949. However, Watson also prompted me to ask for further details of the courses that Plackett and his contemporaries had attended at Cambridge. In response to this request, Plackett (1996) obtained statements from Armitage, Baines, Burman, and Lindley. I shall quote from two of these.

Armitage (1996) notes:

I went to a course on linear algebra by Ingham, in I suppose the second (Prelim.) year. It was one of the clearest courses I attended. Ingham wrote details on the board very clearly, and at a speed that enabled one to copy efficiently, and my notes were consequently very clear. My recollection is that the course covered standard material up to things like orthogonal transformations. Aitken's book [1939a] was recommended reading. There was a parallel course given by someone else based on Farrer's book [presemably Ferrar (1941)].

Lindley's (1996) recollections from the following year were a little different:

I went up to Trinity [College, Cambridge] in October '41, leaving in '43 under wartime regulations.... I do recall having a course on algebra by M. H. A. Newman. It was a chaotic course but nevertheless was one of the best I attended. Unlike Ingham on analysis who was beautifully organised, but terribly dull. When I left I went to do research with Paul Dienes at Birkbeck and made some progress in matrix theory. For example on the square root of a matrix. So Newman must have prepared me well....

Turning to the situation in the U.S.A., we find that Olkin (1995) gives the following account of his mathematical education, in a conference paper:

My introduction to linear algebra at Columbia University [1947–1948] was from the book by Bôcher [1907]. Subsequently, at the University of North Carolina [1948–1951]. I listened to Alfred Brauer and E. T. Browne, who had a set of notes on matrix theory. In the course of study I had occasion to browse through many of the early books on matrix theory: Aitken [1939a]; Turnbull and Aitken [1932]; Frazer, Duncan and Collar [1938]; Schreier and Sperner [1951]; and the extensive compendium by MacDuffee [1933].

In this context Watson (1996) mentions a course of lectures which P. L. Hsu gave in Chapel Hill.

A somewhat fuller account of his own mathematical education is given by Anderson (1995):

In 1940 I was a graduate student in mathematics at Princeton. Since I was interested in statistics, Princeton was a good place for me because of the presence of Sam Wilks.

During the academic year 1940–41, I participated in a seminar reading van der Waerden's *Moderne Algebra* [1931]; however, we concentrated on group theory up to Galois theory, not getting to the part on linear algebra.

In the second semester Wedderburn lectured on Matrices, but he was difficult to follow. My recollection is that only three post-docs stayed with the course. In the same term Halmos lectured on Finite Dimensional Vector Spaces (before the book was published [1942]); of course, he was an excellent expositor and was very popular.

In preparation for taking the Preliminary Oral Subject Matter Exam early in the fall of 1941–2, I studied Bôcher's *Introduction to Higher Algebra* [1907] by myself over the summer. This seemed to be the standard book covering linear transformations and quadratic forms. It included a good deal of algebraic calculation in terms of the components. In the week before Prelims I read what I could of *An Introduction to the Theory of Canonical Matrices* by Turnbull and Aitken [1932].

For my dissertation research Wilks suggested that I look at a couple of papers of Fisher on discriminant analysis [1936, 1938]. That led me to several topics in multivariate analysis including the noncentral Wishart distribution tests of ranks of matrices, and characteristic roots and vectors [of covariance matrices]. Then I had to develop a working knowledge of matrices. My dissertation was written in terms of components, but the methods were matrix methods. (Wilks seemed to be more comfortable in terms of components.)

At the end of World War II, during which time I engaged in wartime "defense" research, I went on to the Cowles Commission for Research in Economics at the University of Chicago. The econometricians (three of whom late received Nobel Prizes) and statisticians (two graduate students and I) devoted ourselves to developing the theory and methods and applications of simultaneous equation models. That research involved heavy matrix calculations. At that time Herman Rubin and I worked on factor analysis as well. By the time I went to Columbia in the fall of 1946, I was completely comfortable with matrices.

Readers are referred to the book edited by Koopmans (1950) for a detailed account of the work performed by the Cowles Commission research workers at this time. Although Turnbull and Aitken's (1932) book outlines several possible applications of matrix algebra in the context of statistical problems, and although several research papers in this area had been published between the years 1935 and 1950, it seems that the first textbooks to introduce matrix methods to statisticians date from the late 1950s and include those of Anderson (1958), Kendall (1957), and Plackett (1960). The books published a few years later by Rao (1965) and Seber (1966) should also be mentioned.

5. DISSEMINATION OF MATRIX THEORY IN ECONOMICS

Despite the crucial role played by matrix theory in the development of statistical methods suitable for application to the simultaneous equations model of econometric theory by Anderson and his colleagues at the Cowles Commission in the late 1940s, the subject was still not well received by mainstream economists in the early 1960s. If these features of the situation had been fully appreciated by Grattan-Guiness (1994, p. 67), he would presumable not have been so surprised to find that Dorfman, Samuelson, and Solow (1958) had felt obliged to include an appendix on elementary matrix algebra in their text on linear programing, as these authors were themselves economists and they were writing a book intended for private study by fellow economists.

As further evidence of this late acquisition of matrix techniques by economists, we note that the first edition of Johnston's (1963) standard textbook of econometric theory was praised for its expository chapter on elementary matrix algebra.

6. THE TERMINOLOGY OF SPECTRAL THEORY

Finally we note that Grattan-Guiness and Ledermann (1994, p. 785) continued their account of matrix theory by criticizing

> The popularization of the appalling non-words "eigenvalue" and "eigenvector" created out of (absurdly partial) translations of the German words *Eigenwert and Eigenvektor*. The properly English phrases "latent root" and "latent vector" have been employed in this article. The former was introduced in Sylvester *1883*....

Although there is some merit in this argument, particularly if one is a native German speaker, it is difficult to take the matter as seriously as the authors appear to do when one realizes that the English language is already replete with such familiar twentieth century hybrid words as "automobile," "multilith," "photocopy," "telecommunication," and "television," and the rather less familiar "audiology" and "criminometrics," whilst recent additions to statistical terminology include nonparametric," "semiparametric," and even "seminonparametric." Whilst it may still be possible to eliminate some of the more obnoxious of these recent coinings, the process has surely gone too far in the case of interest here. The words "eigenvalue" and "eigenvector" are now so widely used by mathematicians, most notably by J. H. Wilkinson in the title of his influential textbook *The Algebraic Eigenvalue Problem* (1965), that their suppression would now seem to be impossible.

I am indebted to Ted Anderson, Peter Armitage, Arnold Baines, Maurice Bartlett, Peter Burman, Ivor Grattan-Guinness, Walter Ledermann, Dennis Lindely, Ingram Olkin, Robin Plackett, and Geoffrey Watson for their valuable help in the preparation of this paper.

REFERENCES

1 Aitken, A. C. 1935. On least squares and linear combination of observations, *Proc, Roy. Soc. Edinburgh Part A* 55:42–47.
2. Aitken, A. C. 1939a. *Determinants and Matrices*, Oliver and Boyd, Edinburgh; reprinted by Greenwood Press, London, 1983.
3. Aitken, A. C. 1939b. *Statistical Mathematics*, Oliver and Boyd, Edinburgh.
4. Aitken, A. C. 1995. *To Catch the Spirit*, Univ. of Otago Press, Dunedin.
5. Anderson, T. W. 1958. *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.
6. Anderson, T. W., 1995. Personal communication.
7. Armitage, P. 1996. Personal communication.
8. Bôcher, 1907. *Introduction to Higher Algebra*, Macmillan, New York.
9. A. L. Cauchy, 1829. Sur l' équation à l' aide de laquelle on détermine les inégalités séculaires des mouvements des planètes, *Exercises Math.* 4:140–160; reprinted in *Oeuvres* (*Ser.* 2) Vol. 9, Gauthier-Villar, Paris, 1891, pp. 174–195.
10. Cayley, A. 1858. Memoir on the theroy of matrices, *Philos. Trans. Roy. Soc. London* 148:17–37 (1858).
11. Courant, R. and Hilbert, D. 1924, 1938. *Methods of Mathematical Physics*, 2 vols., Julius Springer, Berlin; English transl. Interscience, New York, 1953.
12. Cullis, C. E. 1913–1925. *Matrices and Determinoids*, 3 vols., Cambridge U. P. Cambridge.
13. Dorfman, R., Samuelson, P. A., and Solow, R. M. 1958. *Linear Programming and Economic Analysis*, McGraw-Hill, New York.
14. Eisenstein, G. 1844. Algemeine Untersuchungen über die Formen dritten Grades mit drei Variabeln, welche der Kreistheilung ihre Entstehung verdanken, *J. Reine Angew. Math.* 28:289–374.
15. Ferrar, W. L. 1941. *Algebra: A Text-book of Determinants, Matrices and Algebraic Forms*, Oxford U. P., Oxford, 1941.
16. Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems, *Ann. Eugen.* 7:179–188.
17. Fisher, R. A. 1938. The statistical utilisation of multiple measurements, *Ann. Eugen.* 8:376–386.
18. Frazer, R. A., Duncan, W. J., and Collar, A. R. 1938. *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, MacMillan, London; reprinted by AMS Press, New York, 1982.
19. Frobenius, G. 1877. Ueber das Pfaffsche Problem, *J. Reine Angew. Math.* 82:230–315.
20. Gauss, C. F. 1801. *Disquisitiones Arithmeticae*, Gerhardt Fleischer; reprinted in his *Werke*, Vol. 1, Univ. of Göttingen, Göttingen, 1870.
21. Grattan-Guinness, I. 1994. "A new type of question". On the prehistory of linear an nonlinear programming, Hist. Modern Math. 3:43–89.
22. Grattan-Guinness, I. and Ledermann, W. 1994. Matrix theory, in (I. Grattan-Guiness, Ed.) *Companion Encyclopedia to the History and Philosophy of the Mathematical Sciences* Routledge, London, pp. 775–786.

23. Halmos, P. R. 1942. *Finite-Dimensional Vector Spaces*, Princeton U. P., Princeton, N.J.
24. Hawkins, T. W. 1975. The theory of matrices in the 19th century, in *Proceedings of the International Congress of Mathematics*, Vancouver, Canada, Vol. 2, pp. 561–570.
25. Hermite, C. 1854. Sur la Théorie des formes quadratiques ternaires indéfines, *J. Reine Angew. Math.* 47; reprinted in his *Oeuvres* vol. 1, Gauthier-Villars, Paris, 1905, pp. 193–199.
26. Jacobi, C. G. J. 1827. Über die Pfaffsche Methode, eine gewöhnliche lineare Differentialgleichung zwischen $2n$ Variabeln durch ein System von n Gleichungen zu integrieren, *J. Angew. Math.* 2:347–357; reprinted in his *Werke*, Vo. 4, Reimer, Berlin, 1886, pp. 17–29.
27. Johnston, J. 1963. *Econometric Methods*, McGraw-Hill, New York.
28. Kendall, M. G. 1957. *A Course in Multivariate Analysis*, Griffin, London.
29. Koopmans, T. C. 1950. *Statistical Inference in Dynamic Economic Models*, Wiley, New York.
30. Lagrange, J. L. 1775. Nouvelle solution du problème du rotation d'un corps de figure quelconque qui n'est animé par acune force accélératrice, *Nouv. Mem. Acad. Berlin*; reprinted in his *Oeuvres*, Vol. 3, Gauthier-Villars, Paris, 1869, pp. 577–616.
31. Ledermann, W. W. 1968. A. C. Aitken's work in pure mathematics, *Proc. Edinburgh Math. Soc. Ser. 2* 16;162–165.
32. Lindley, D. V. 1996. Personal communication.
33. MacDuffee, C. C. 1933. *The Theory of Matrices*, Julius Springer, Berlin; reprinted by Chelsea, New York.
34. Mirsky, L. A. 1955. *An Introduction to Linear Algebra*, Oxford U. P., Oxford.
35. Olkin, I. 1955. A. C. Aitken on compounds, presented at the Fourth International Workshop on Matrix Methods for Statistics, Montreal, Canada.
36. Plackett, R. L. 1960. *Principles of Regression Analysis*, Oxford U. P., Oxford.
37. Plackett, R. L. 1996. Personal communications.
38. Rao, C. R. 1965. *Linear Statistical Inference and Its Applications*, Wiley, New York.
39. Schreier, O., and Sperner, E. 1951. *Introduction to Modern Algebra and Matrix Theory*, Chelsea, New York.
40. Seber, G. A. F. 1966. *The Linear Hypothesis: A General Theory*, Griffin, London.
41. Sturm, C. 1829. Extrait d'un mémoire sur l'intégration d' un système d' équations différentielles linéaires, *Bull. Sci. Math.* 12:313–322.
42. Sylvester, J. J. 1850. Additions to the articles, "On a new class of theorems," and "On Pascal's theorem," *Philos. Mag. Ser. 3* 37:363–379; reprinted in his *Mathematical Papers*, Vol. 1, Cambridge U. P., Cambridge, 1904, pp. 145–151.
43. Sylvester, J. J. 1851. On the relation between the minor determinants of linearly equivalent quadratic forms, *Philos. Mag. Ser. 4* 1:295–305; reprinted in his *Mathematical Papers*, Vol. 1, Cambridge U. P., Cambridge, 1904, pp. 241–249.
44. Sylvester, J. J. 1883. On the equation to the secular inequalities in the planetary theory, *Phil. Mag. Ser. 5* 16:267–269.

45. Tee, G. 1980. Two New Zealand Mathematicians, in *First Australian Conference on the History of Mathematics*, Monash U., Clayton, Australia, pp. 182–199.
46. Turnbull, H. W., and Aitken, A. C. 1932. *An Introduction to the Theory of Canonical Matrices*, Blackie, London; reprinted by Dover, New York, 1961.
47. van der Waerden, B. L. 1931. *Moderne Algebra* J. Springer, Berlin.
48. Watson, G. S. 1995, 1996. Personal communications.
49. Wedderburn, J. H. M. 1934. *Lectures on Matrices*, AMS Press, New York; reprinted by AMS Press, Providence, 1960, and Dover, New York, 1964.
50. Weierstrass., K. 1868. Aur Theorie der bilinearen und quadratischen Formen, *Monatsber. Berlin. Akad. Wiss.* reprinted in his *Mathematische Werke*, Vol. 2, Mayer und Muller, Berlin, 1895 pp. 19–44.
51. Whittaker, E. T., and Robinson, G. 1924. *The Calculus of Observations*, Blackie, London; 4th ed. (1944) reprinted by Dover, New York, 1967.
52. Whittaker, J. M. 1968. Alexander Craig Aitken, 1895–1967, *Biograph. Mem. Fellows Soc. London* 14:1–8.
53. Wilkinson, J. H. 1965. *The Algebraic Ekgenvalue Problem*, Oxford U. P., Oxford, 1965.

Received 16 December 1995; final manuscript accepted 18 May 1996